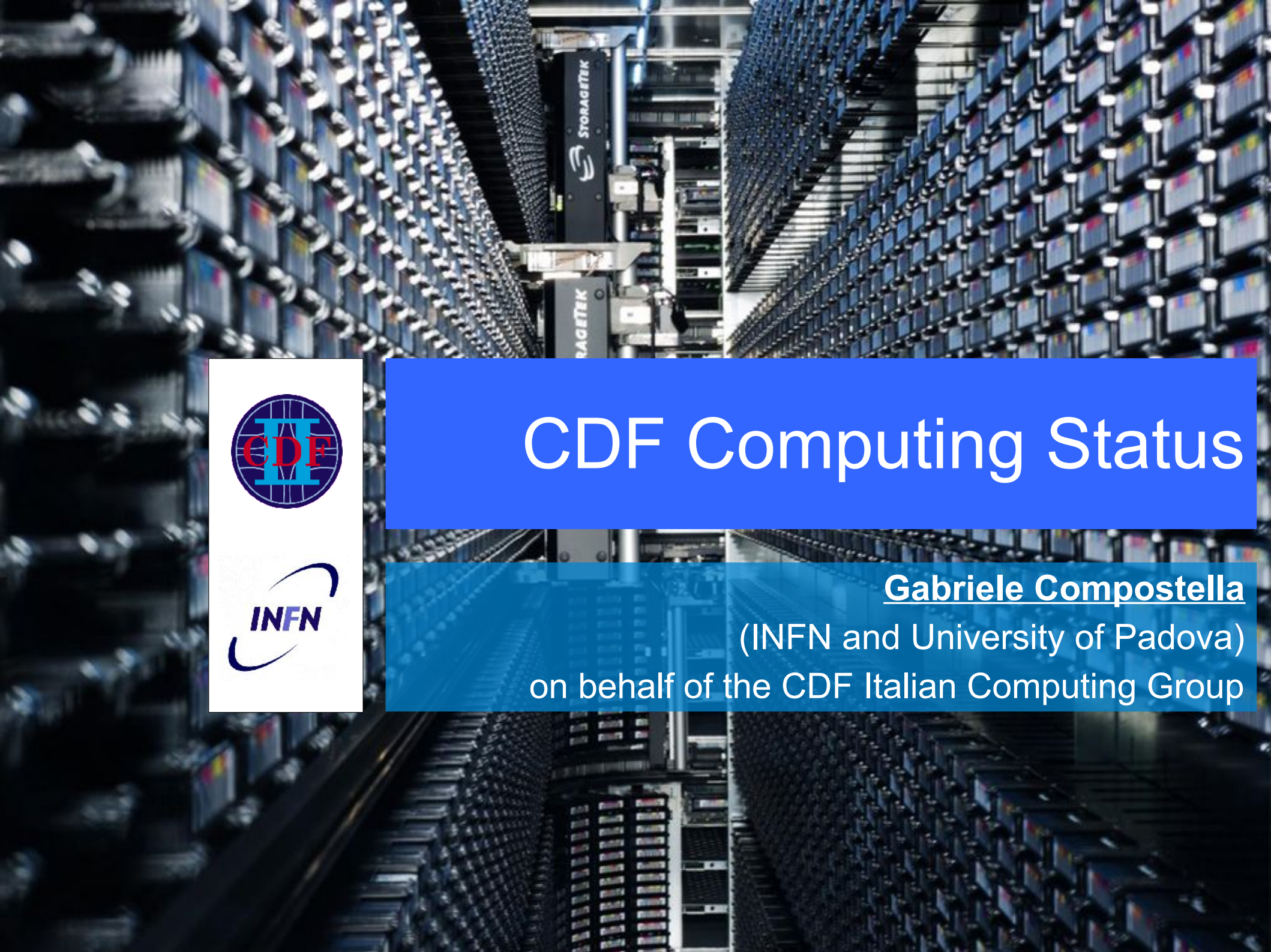


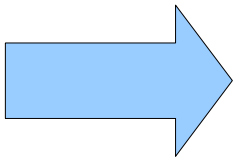
CDF Computing Status

Gabriele Compostella
(INFN and University of Padova)
on behalf of the CDF Italian Computing Group

CDF Computing will be needed for a while...

- *Offline systems are stable*
- *Need to keep up with GRID development*
- *Keep supporting systems and users: data production, Monte Carlo production, ntupling and few applications*

- *Recently moved every computing resource in the lab (desktops, servers and on-site farm worker nodes) to Scientific Linux 5*
- *Performed a full compatibility test of the experiment's code, developed optimized versions of the code for SL5 and guidelines for offsite installations*



This should ensure sustainability of CDF offline computing and smooth operations over the next years, as long as people and resources are available
(i.e. PAC documents foresee ~15 FTE for the offline only)

Data production summary

Data delivery for some recent periods:

Data taking period	lumi	Calib ready	Prod start	Prod done	Ntuples
02/26/10 - 04/13/10	360 pb-1	05/24/10	06/07/10	07/15/10	08/02/10
01/06/10 - 02/25/10	333 pb-1	03/31/10	03/31/10	04/16/10	05/03/10
10/25/09 – 01/05/10	422 pb-1	02/25/10	02/25/10	03/30/10	04/07/10
09/15/09 – 10/25/09	189 pb-1	12/14/09	12/14/09	01/05/10	01/20/10

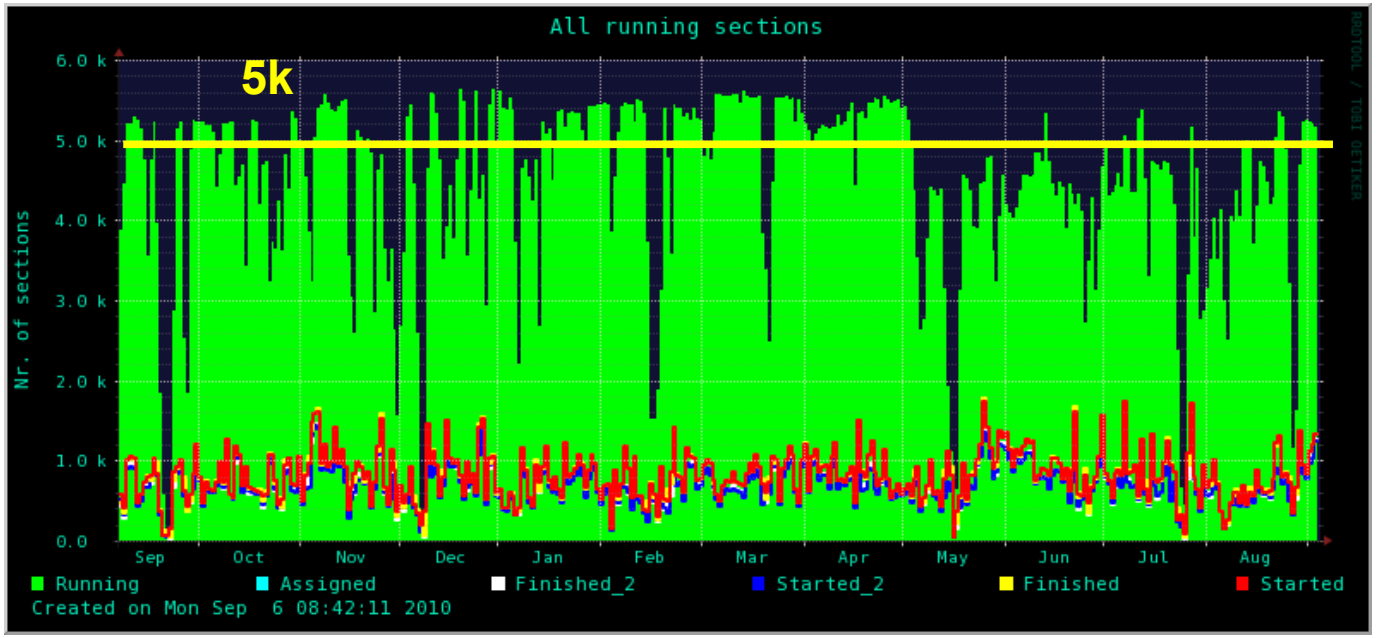
Production running smooth, ~12 weeks to have data available for user analysis

Fermilab onsite computing

Job submission:

- *currently 2 main portals are available in the US and 2 main portals in Europe*
- *in both cases one portal accesses dedicated resources, the other one GRID resources*
- *Resources available in Asia are being merged into US portals*

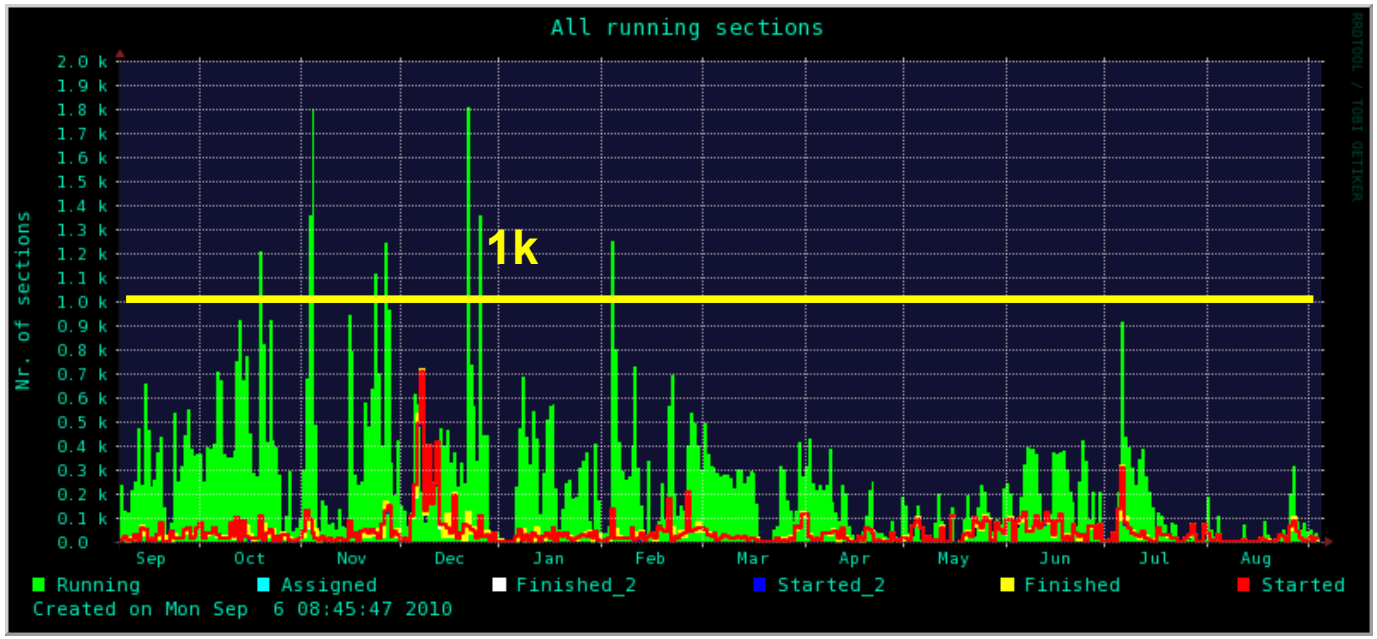
Onsite computing resources are accessed through a single job submission portal and are used mainly for data processing and user analysis:



CDFGRID Portal

Onsite resources are saturated

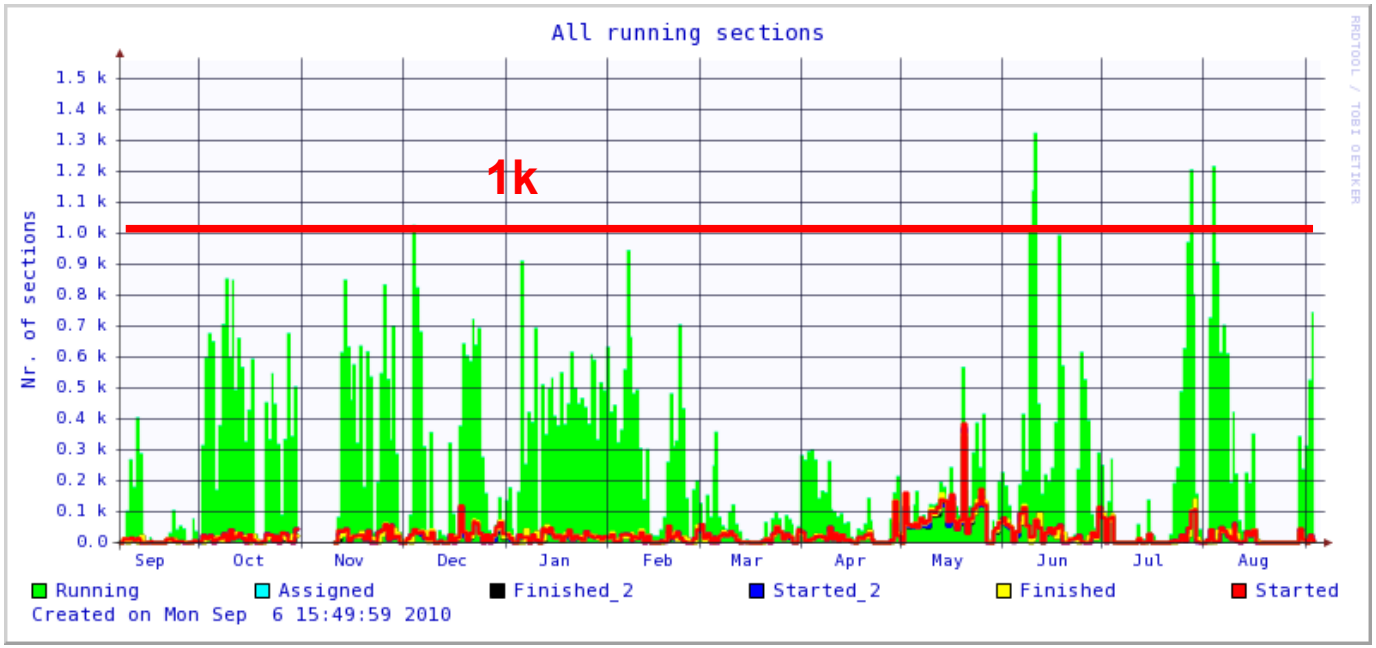
An additional portal is used to access computing resources in the Open Science Grid in an opportunistic way



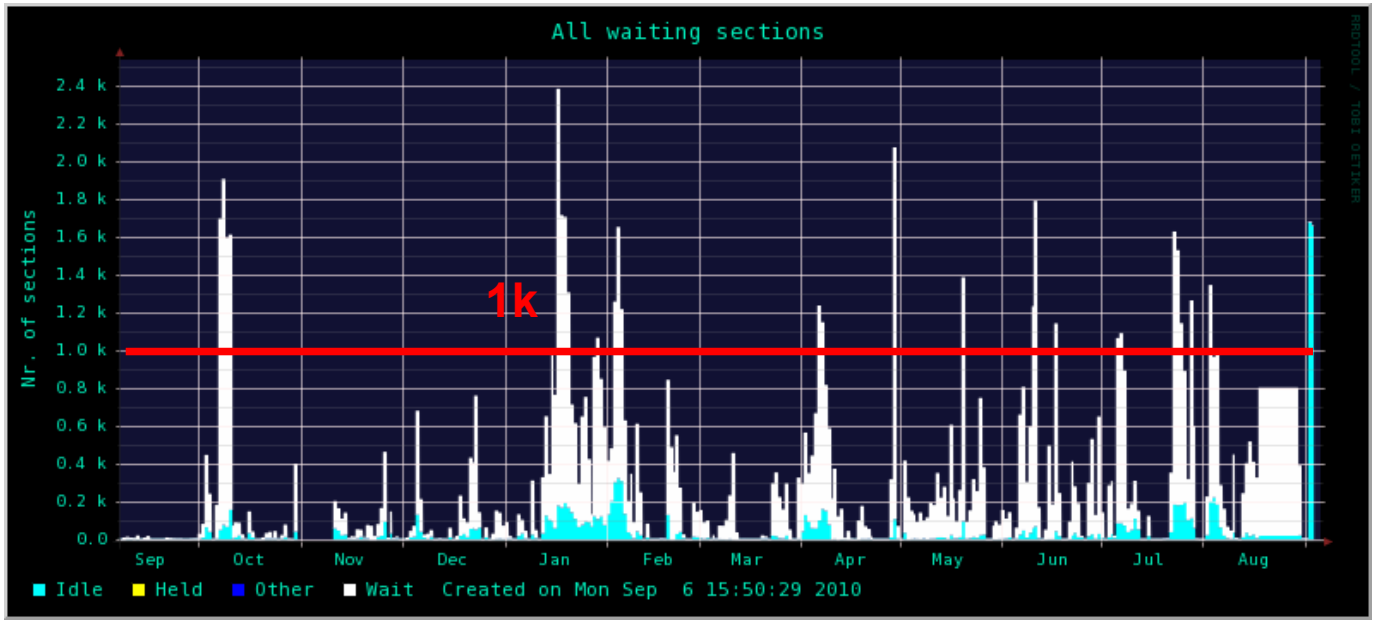
NamGRID Portal

- *Mainly used for MC production, no data access*
- *Needs constant interaction with site admins*
- *Resources are not guaranteed to be available for CDF*

In Italy CNAFCAF portal is used to submit jobs to the Tier1:

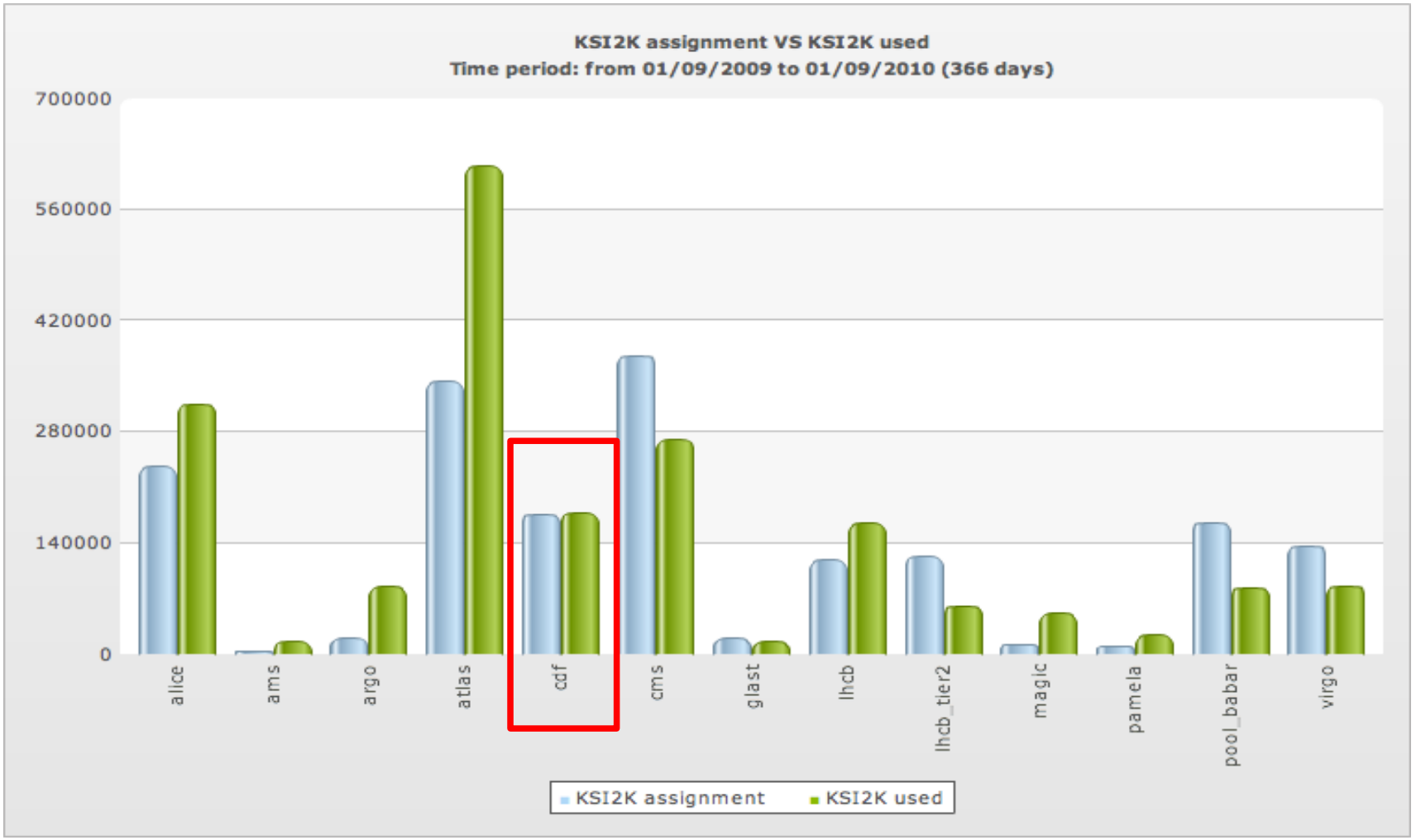


CNAFCAF Portal



Used for MC production and user data analysis: a subsample of CDF data interesting for the italian collaboration is stored at CNAF for efficient data analysis

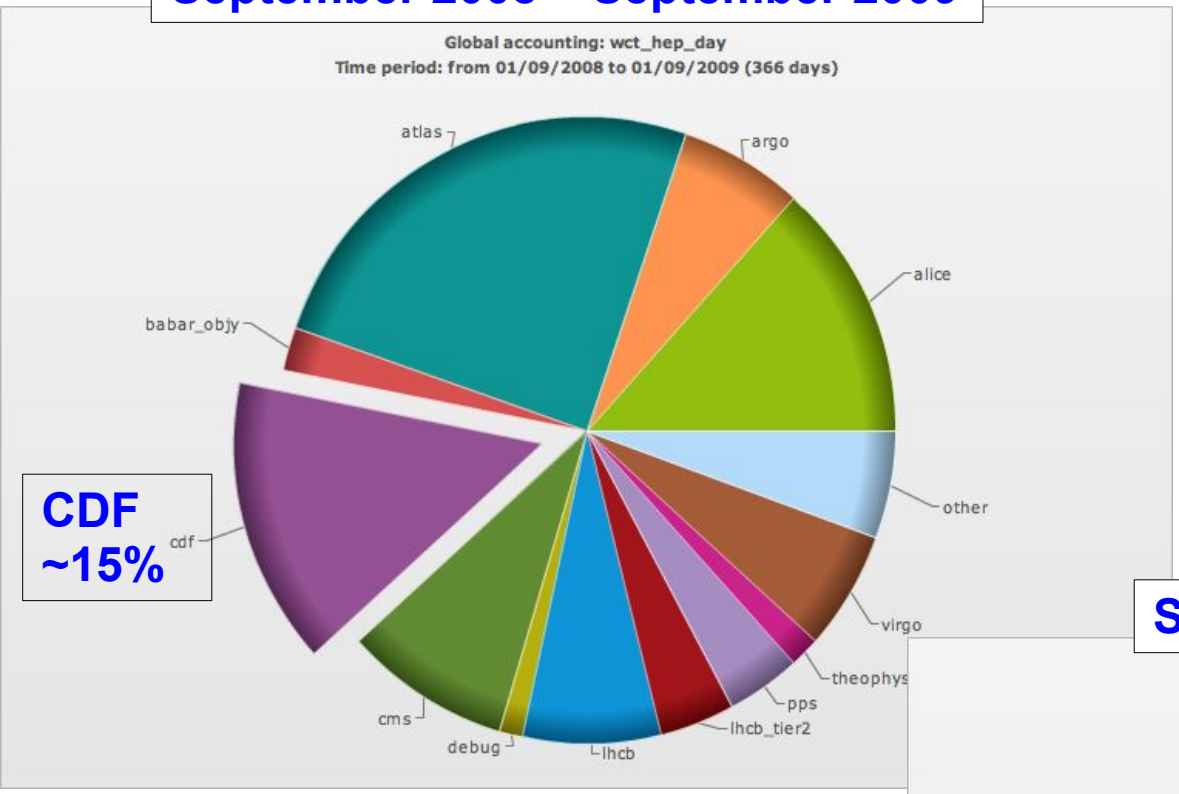
Resource usage at CNAF-Tier1



During September 2009 – September 2010 CDF used its full assignment at Tier1

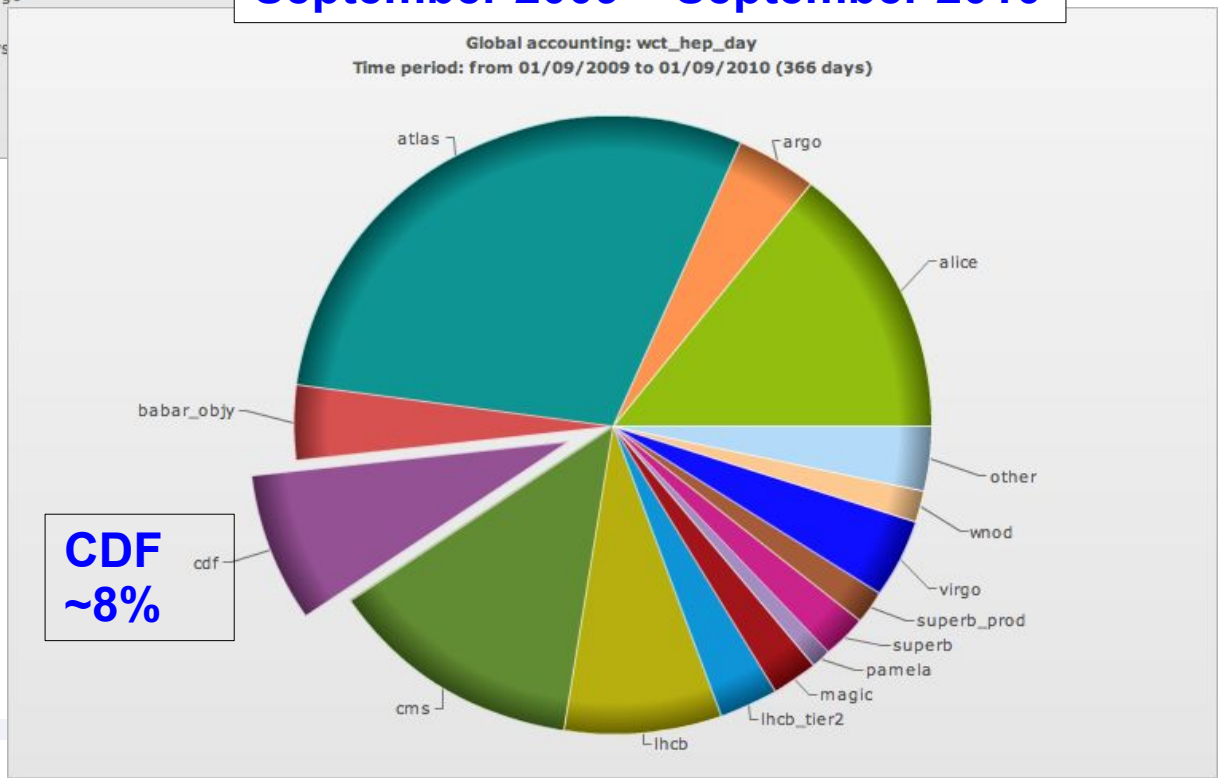
LHC experiments and CDF @T1

September 2008 – September 2009



CDF ~15%

September 2009 – September 2010



CDF ~8%

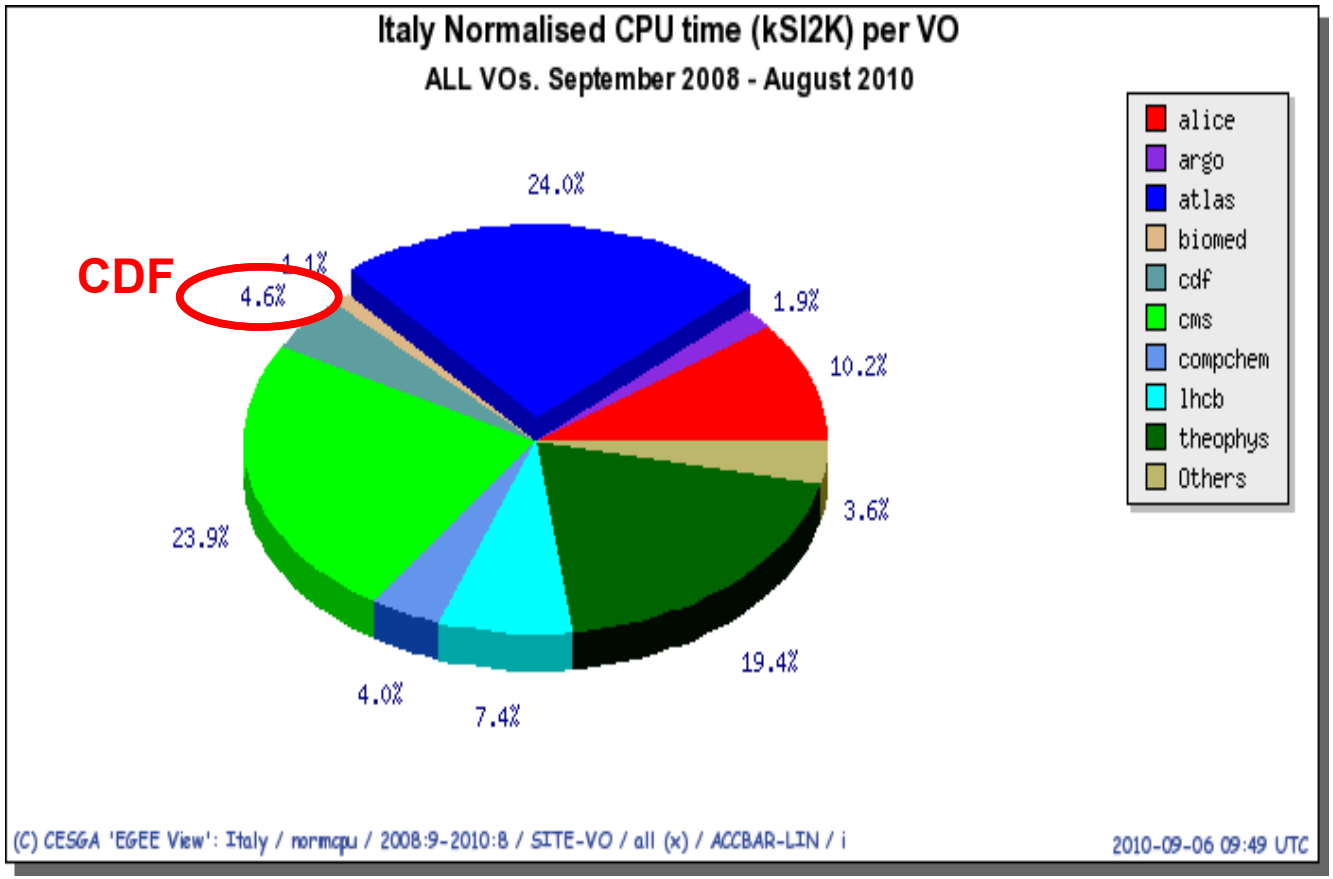
Lost some resources due to the ramp up in the activity of LHC experiments

Grid Computing in Italy and Europe: LcgCAF

Get additional resources exploiting Italian and European Grid sites:

Site	Country
CNAF-T1	Italy
INFN-Padova	Italy
INFN-Bari	Italy
INFN-Legnaro	Italy
INFN-Roma1	Italy
INFN-Roma2	Italy
INFN-Catania	Italy
INFN-Pisa	Italy
FZK-LCG2	Germany
IN2P3-CC	France
IEPSAS	Slovakia
IFAE	Spain
PIC	Spain

- *Mainly used for MC production*
- *Had some problems with the SL5 transition*
- *Usage shows spikes of several thousands, the average is around few hundreds*



Overall computing activity on italian sites in the past 2 years

Available Disk Resources

At Fermilab:

- *Temporary user manageable storage is available*
- *Data stored in dcache managed tapes*
- *Dcache based diskpool for reduced analyses datasets, managed by physics groups*
- *General pourpose data servers for physics groups and operations*

No Grid technology available!

At CNAF:

- *Storage Element (managed with STORM) at CNAF: buffer for Monte Carlo production and analysis space for Italian users*
- *Temporary storage for imported CDF data*
- *Several CDF datasets have a copy at CNAF to allow direct data access and analysis*

Italian resources in Fermilab:

- *Want to use GRID-aware technologies to allow an easy way to transfer analysis datasets back and forth between Italy and FNAL (i.e. Padova SRM located at FNAL, works and could be replicated for other users)*

Data Reprocessing:

- *Around half of the data collected up to now will soon be reprocessed to allow a better tracking and b-tagging efficiency*
- *This will require a re-ntuplization of datasets used for B and high-Pt analyses*
- *In the meanwhile, data production will progress as new data are collected*
- *In addition to this, the budget of the Computing Division for CDF computing has been cut down to a half with respect to previous year*

→ *This will likely result in a higher than average load on the computing resources located off-site, like CNAF and LCGCAF*
→ *We plan to take part of the reprocessing at CNAF*

Development:

- *For easier and more efficient management, we would like to merge all European resources and CNAF under a single portal and do the development needed to make the current systems more stable*

- *Tevatron program has been approved for FY 2011 and beyond*
- *This was not taken into account in the early Tier1 Plan for 2011*

Richieste:

DISK:	30 TB	(had 270 TB in 2010)
CPU:	7000 HS06	(had 6200 HS06 in 2010)

Need to come up with a plan for CDF computing for the next years, in case of RUNIII !